DOCUMENT RESUME

ED 069 672 TM 002 123

AUTHOR Veldman, Donald J.
TITLE Automated Sentence Completion Scoring.
NOTE 5p.; Presented to American Psychological Association

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Affective Tests; College Students; *Correlation; *Electronic Data Processing; Factor Analysis; Input Output Devices; Item Analysis; *Measurement Techniques; Norms; Personality Tests; Questionnaires; *Research Methodology; Sex Differences; Teacher Education; Test Interpretation; Test Reliability

ABSTRACT

A 62-item form of the sentence-completion technique requiring one-word responses was administered to 1718 undergraduates in teacher education. The data were punched on cards and lists of different responses were compiled. Responses indicating evasion, hostility, anxiety and depression were identified for each stem to form a scoring "dictionary." A computer program scored all protocols for these four variables, as well as populars, average response length and repetitions. Descriptive statistics, internal consistency reliabilities, and intercorrelations of the variables are reported, as well as concurrent validities against a self-report questionnaire and comparisons of sex and teaching-level subsamples. (Author)

ED 069672

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Automated Sentence Completion Scoring

Donald J. Veldman
The University of Texas at Austin

A 62-item form of the sentence-completion technique requiring one-word responses was administered to 1718 undergraduates in teacher education. The data were punched on cards and lists of different responses were compiled. Responses indicating evasion, hostility, anxiety and depression were identified for each stem to form a scoring "dictionary." A computer program scored all protocols for these four variables, as well as populars, average response length and repetitions. Descriptive statistics, internal consistency reliabilities, and intercorrelations of the variables are reported, as well as concurrent validities against a self-report questionnaire and comparisons of sex and teaching-level subsamples.

TM 002 123

FILMED FROM BEST AVAILABLE COPY

AUTOMATED SENTENCE COMPLETION SCORING

Donald J. Veldman

The University of Texas at Austin

Personality assessment techniques vary considerably in terms of the degree of structure provided to the respondent. Questionnaires require only choices among supplied alternatives, while story-telling tasks may provide no structured stimuli at all. The sentence completion technique lies somewhere between these two extremes, and is a popular technique for personality assessment.

Interpretation of verbal responses to sentence completion items is usually a matter of clinical intuition, the extreme variety of possible responses makes development of objective scoring procedures very difficult. To facilitate the construction of a computer-based scoring system for sentence completion responses Veldman, Menaker, and Peck (1969) devised a form with instructions to use no more than one word to complete each stem. A great deal of experimentation with various types of stems and response-reduction techniques led to the current system, which is presently in use at The University of Texas as part of a personal assessment program for all students in the undergraduate teacher education program.

Development of the System

A total of 1718 protocols were gathered using a 62-item form during a full year in the College of Education. All responses were transferred to punch cards. A ten-character response limit permitted data from one respondent to be punched in nine cards. A computer routine was then used to compile for each stem a list of all different responses and their frequencies. These lists varied in length from 33 to 538 responses.

The author then examined each list to identify five types of responses. Populars were defined as words used by more than ten percent of the respondents. Evasive responses included blanks, proper names, stem word repetitions and other cryptic answers. Hostile responses were primarily instances of derogation of other people. Anxiety indicators included both direct and projected expressions of fear and tension. Depression indicators included direct and projected self-derogation and pessimism.

The E, H, A, and D response words were converted to root forms (confused, confusing, confusion = confus) and punched on cards to assemble a scoring "dictionary." No response word was coded for more than one category on one stem, although the same word might be coded differently on different stems.

A scoring program was then written and applied to all of the original 1718 protocols. In addition to the four "clinical" scores, the program computed the average length in characters of the subjects' 62 responses, and also counted the number of repetitions of response words

within the protocol. An eighth total negative affect score was computed as the sum of hostility, anxiety, and depression scores.

Statistical Properties of the Scores

Central tendency and variability indices and alpha coefficients of internal consistency were computed for all eight variables.

Average Response Length (ARL). Mean = 6.60, sigma = .51, alpha = .77.
Repetitions (REP). Mean = 1.0, sigma = 4.2, alpha = .62.
Populars (POP). Mean = 19.0, sigma = 4.7, alpha = .59.
Evasion (EVA). Median = 1.5, alpha = .70.
Hostility (HOS). Median = 1.5, alpha = .49.
Anxiety (ANX). Median = 1.5, alpha = .33.
Depression (DEP). Median = 1.5, alpha = .50.
Total Negative Affect (TNA). Mean = 6.3, sigma = 3.8, alpha = .66.

The evasion, hostility, anxiety, and depression score distributions were quite similar. Roughly 25% of the respondents gave zero, one, two, and more than two of each.

Intercorrelations of the scores are shown in Table 1. Short responses tend to be repeated and popular. Populars tend to be inversely related to evasion, hostility, and depression.

Concurrent Validity

Another instrument in the personal assessment battery completed by these students was the 56-item Adjective Self-Description (Veldman and Parker, 1970), which yields seven trait measures. Correlations between these and the sentence completion variables are shown in Table 2.

Response length and repetitions had no strong correlates. Populars were associated with warmth, lack of abrasiveness, efficiency, and practicality. Evasion had no strong correlates. Hostility was related to abrasiveness, lack of warmth, and anxiety. Anxiety correlated with its counterpart. Depression was related to abrasiveness, inefficiency, and anxiety.

The fact that these relationships were only moderate on the whole can be interpreted as evidence that the sentence completion method is tapping rather different sources of variance than the highly structured questionnaire. Such lack of redundancy is a desirable feature for use in a multi-instrument battery.

Comparison of Sexes and Teaching Levels

The total group of 1718 students contained 341 males and 1377 females. Among the females were 457 elementary and 920 secondary majors. In Table 3 are shown the results of analyses of variance comparing these groups.

Among the female students, the elementary majors use more popular responses and show less evasion, hostility, and depression. Among the

secondary majors, males give shorter responses and more repetitions, use fewer populars, and more evasives, indicate greater hostility and less anxiety than do the females.

Applications of the System

A major advantage of the one-word sentence completion technique is that the same protocols may serve multiple purposes in a large-scale assessment program. The computer-generated scores can be automatically scanned to identify students who reveal clear evidence of need for personal counseling. When such cases are identified by such a screening system, or when students voluntarily seek counseling services, the original protocols are also available for clinical interpretation along with other assessment data derived from more and from less structured stimulus instruments.

To aid clinical use of the instrument, a computer program has been developed which prints a one-page summary of the results of the scoring system, including verbatim reproduction of E, H, A, and D responses embedded in the sentences in which they were given.

REFERENCES

Veldman, D. J., Menaker, S. L., & Peck, R. F. Computer scoring of sentence completion data. Behavioral Science, 1969, 14, 501-507.

Veldman, D. J., & Parker, G. V. C. Adjective rating scales for self description. Multivariate Behavioral Research, 1970, 5, 295-302.

Table I. Intercorrelations of the Sentence Completion Variables

Scale	ARL	REP	POP	EVA	HOS	ANX	DEP	TNA
ARL	--	-46	-27	-08	-01	06	-08	-03
REP	-46	--	36	-06	-16	-10	-06	-16
POP	-27	36	--	-25	-30	-08	-27	-33
EVA	-08	-06	-25	--	-02	-10	-04	-08
HOS	-01	-16	-30	-02	--	13	32	73
ANX	06	-10	-08	-10	13	--	20	57
DEP	-08	-06	-27	-04	32	20	--	76
TNA	-03	-16	-33	-08	73	57	76	--

Table 2. Correlations of Sentence Completion with ASD Variables

ASD Variables	ARL	REP	POP	EVA	HOS	ANX	DEP	TNA
Social Warmth	-02	04	18	-07	-21	-01	-13	-18
Social Abrasiveness	-03	-07	-20	08	23	06	25	27
Ego Organization	03	06	15	-06	-18	-07	-24	-24
Introversion/Extraversion	-07	03	-05	-01	01	12	17	14
Neurotic Anxiety	00	-03	-09	-03	20	25	30	36
Idealism	11	-09	-17	10	09	-05	05	05
Social Attractiveness	08	-01	07	-08	-02	-09	-15	-13

Table 3. Mean Differences Between Sexes and Teaching Levels

OWSC Variable	FE Mean	P level	FS Mean	P level	MS Mean
Average Response Length	6.64	ns	6.62	.0006	6.50
Repetitions	9.60	ns	9.87	.0004	10.87
Populars	20.11	.005	19.35	<.0001	16.71
Evasion	1.61	.002	2.03	<.0001	2.87
Hostility	1.83	.002	2.17	.0002	2.69
Anxiety	2.04	ns	1.96	<.0001	1.37
Depression	2.04	.09	2.23	ns	2.40
Total Negative Affect	5.91	.03	6.37	ns	6.47
N	457		920		341